

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-58188

(P2003-58188A)

(43) 公開日 平成15年2月28日 (2003.2.28)

(51) IntCl. <sup>7</sup>	識別記号	F I	テマコード* (参考)
G 1 0 L 15/22		G 0 6 F 3/16	3 2 0 H 5 D 0 1 5
G 0 6 F 3/16	3 2 0		3 4 0 M
	3 4 0	G 1 0 L 3/00	5 7 1 U
G 1 0 L 15/00			5 5 1 Q
15/06			5 7 1 V

審査請求 未請求 請求項の数 7 O L (全 10 頁) 最終頁に続く

(21) 出願番号 特願2001-245707 (P2001-245707)

(22) 出願日 平成13年8月13日 (2001.8.13)

(71) 出願人 000237592

富士通テン株式会社

兵庫県神戸市兵庫区御所通1丁目2番28号

(72) 発明者 中村 英樹

兵庫県神戸市兵庫区御所通1丁目2番28号

富士通テン株式会社内

(74) 代理人 100077517

弁理士 石田 敬 (外4名)

Fターム(参考) 5D015 AA04 BB01 HH13 HH23 LL10

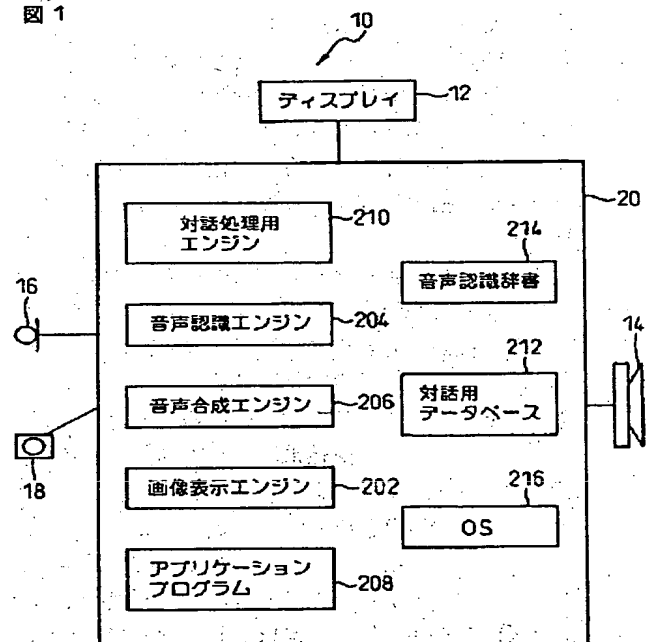
(54) 【発明の名称】 音声対話システム

(57) 【要約】

【課題】 音声認識に当たって認識率の高い音声対話システムを提供する。

【解決手段】 本音声対話システムは、ユーザの発声を認識するための音声認識エンジンと、その認識結果に応じてユーザへの発話を作成する対話処理エンジンと、作成された発話を音声に合成するための音声合成エンジンと、音声認識のための音声パターンを含む辞書データを格納する音声認識辞書と、音声の入出力手段とを備える。さらにこの音声認識エンジンは、前回のユーザ発声の認識結果を音声認識辞書において探索し、探索された認識結果に関連する辞書データを抽出して自身の記憶領域内に設定するよう構成されている。従って、次のユーザの発声の音声認識に当たっては、前回の認識結果に関連する辞書データが使用されるので、認識率が向上する。

図 1



## 【特許請求の範囲】

【請求項1】 ユーザの発声を認識するための音声認識エンジンと、その認識結果に応じてユーザへの発話を作成する対話処理エンジンと、作成された発話を音声に合成するための音声合成エンジンと、前記音声認識のための音声パターンを含む辞書データを格納する音声認識辞書と、音声の入出力手段とを備え、前記音声認識エンジンは前回のユーザ発声の認識結果を前記音声認識辞書において探索し、探索された認識結果に関連する辞書データを抽出して前記音声認識エンジンに設定し次のユーザ発声の音声認識を行なうものである、音声対話システム。

【請求項2】 請求項1に記載の音声対話システムにおいて、前記音声認識エンジンは前回のユーザ発声の認識結果が複数ある場合、そのいずれをも前記音声認識辞書において探索し、いずれの認識結果にも関連する辞書データを前記音声認識エンジンに設定するものである、音声対話システム。

【請求項3】 請求項1に記載の音声対話システムにおいて、前記音声認識エンジンは認識結果の探索対象とする単語属性のリストを有している、音声対話システム。

【請求項4】 請求項3に記載の音声対話システムにおいて、前記音声認識エンジンは前回のユーザ発声の認識結果中に前記リストに対応する単語属性がある場合、それを保存しておくものである、音声対話システム。

【請求項5】 ユーザの発声を認識するための音声認識エンジンと、その認識結果に応じてユーザへの発話を作成する対話処理エンジンと、作成された発話を音声に合成するための音声合成エンジンと、前記音声認識のための音声パターンを含む辞書データを格納する音声認識辞書と、音声の入出力手段とを備え、前記音声認識辞書は複数の種類の辞書を有し、前記音声認識エンジンは対話状態と該対話状態で使用する種類の辞書の対応テーブルを有し、各対話状態のユーザ発声の認識に当たって前記対応テーブルに従って対応する種類の辞書を前記音声認識辞書から抽出して使用するものである、音声対話システム。

【請求項6】 請求項5に記載の音声対話システムにおいて、前記音声認識エンジンは前記対応テーブルに従って複数の種類の辞書を前記音声認識辞書より抽出し使用するものである、音声対話システム。

【請求項7】 ユーザの発声を認識するための音声認識エンジンと、その認識結果に応じてユーザへの発話を作成する対話処理エンジンと、作成された発話を音声に合成するための音声合成エンジンと、前記音声認識のための音声パターンを含む辞書データを格納する音声認識辞書と、音声の入出力手段とを備え、前記音声認識エンジンは所定の対話タスクを達成する複数の単語属性中で前回のユーザ発声の認識時に使用された単語属性以外の単語属性を含む辞書データを前記音声認識辞書から抽出し

て次回のユーザ発声の認識に使用するものである、音声対話システム。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、各種の機器例えば車両に搭載されたカーナビゲーションシステム、オーディオ機器などの操作を、ユーザとの対話に基づいて自動的に遂行する音声対話システムに関する。

## 【0002】

【従来の技術】カーナビゲーションシステム、オーディオ機器など操作が複雑な機器では、機器を操作するために音声対話システムが用いられる。このシステムでは、予め決められたステップに基づいて作成される質問をシステムが発話し、この発話に対するユーザの回答を音声認識し、認識結果に基づいて新たな発話を作成するという手順によって、機器の操作に必要な情報をユーザより得ている。

## 【0003】

【発明が解決しようとする課題】ところがユーザの発声は、発声の速度、音質、イントネーションなどが種々万別であり、従って音声パターンの比較照合による音声認識では、往々にして認識エラーが発生する。このようなエラーが発生すると、ユーザはそのエラーを訂正するための発声を行なった後、同じ発声を繰り返す必要があり、対話の進行速度が大幅に低下する。従って、認識エラーが度々発生すると、機器の操作性は大幅に低下する。

【0004】本発明は、従来の音声対話システムの上記の点に関してなされたもので、認識率が高く、従って操作性の良い音声対話システムを提供することを目的とする。

## 【0005】

【課題を解決するための手段】上記課題を解決するために、本発明の第1の態様では、ユーザの発声を認識するための音声認識エンジンと、その認識結果に応じてユーザへの発話を作成する対話処理エンジンと、作成された発話を音声に合成するための音声合成エンジンと、前記音声認識のための音声パターンを含む辞書データを格納する音声認識辞書と、音声の入出力手段とを備え、前記音声認識エンジンは前回のユーザ発声の認識結果を前記音声認識辞書において探索し、探索された認識結果に関連する辞書データを抽出して該音声認識エンジンに設定し次のユーザ発声の音声認識を行なう、音声対話システムが提供される。

【0006】この音声対話システムでは、音声認識エンジンは、前回の認識結果を音声認識辞書内で探索し、探索された認識結果に関連する辞書データのみを音声認識辞書から抽出し、これを自身の記憶領域内に辞書として設定する。次回の音声認識は、音声認識エンジンにおいてこの様にして設定された辞書を用いて実施される。そ

のため、前回の認識結果に関連しない認識結果を得ると言う認識エラーの発生が防止される。また、認識のための辞書データが大幅に絞り込まれるので、認識速度が向上し、しいては対話の応答速度が向上する。

【0007】上記態様の音声システムにおいて、前回のユーザ発声の認識結果が複数ある場合、音声認識エンジンはそのいずれをも前記音声認識辞書において探索し、いずれの認識結果にも関連する辞書データを前記音声認識エンジンに設定する。

【0008】これによって、音声認識エンジンに辞書として組み込まれるデータがさらに絞り込まれ、認識エラーの発生率が低下すると共に、対話の応答速度が向上する。

【0009】上記態様の音声システムにおいて、音声認識エンジンは認識結果の探索対象とする単語属性のリストを有している。従って前回の音声認識に当たってこのリストに無い単語が認識された場合は、その単語を除外してリストにある単語のみを辞書中で探索する。これによって、探索対象とする必要がない単語がユーザ発声中に含まれていた場合でも、対処することができる。

【0010】上記態様の音声システムにおいて、前記音声認識エンジンは前回のユーザ発声の認識結果中に前記リストに対応する単語属性がある場合、それを保存しておく。これによって、次のユーザ発声の認識にエラーが生じた場合、保存してある単語属性を利用して辞書の探索、関連する辞書データの抽出を行なうことができる。この結果、認識エラーに対して速やかに対処可能となる。

【0011】本発明の第2の態様では、ユーザの発声を認識するための音声認識エンジンと、その認識結果に応じてユーザへの発話を作成する対話処理エンジンと、作成された発話を音声に合成するための音声合成エンジンと、前記音声認識のための音声パターンを含む辞書データを格納する音声認識辞書と、音声の入出力手段とを備え、前記音声認識辞書は複数の種類の辞書を有し、前記音声認識エンジンは対話状態と該対話状態で使用する種類の辞書の対応テーブルを有し、各対話状態のユーザ発声の認識に当たって前記対応テーブルに従って対応する種類の辞書を前記音声認識辞書から抽出して使用する、音声対話システムが提供される。

【0012】また、この音声対話システムにおいて、前記音声認識エンジンは前記対応テーブルに従って複数の種類の辞書を前記音声認識辞書より抽出し使用する。

【0013】この音声対話システムによれば、音声認識エンジンは対話状態に応じて予め設定されている対応テーブルに従って、必要とする種類の辞書のみを音声認識辞書から抽出し、これを自身の記憶領域に設定する。従って、音声認識に当たって、必要とする辞書以外は認識作業の対象とならないので、音声認識エンジンにかかる負荷が低減され、その結果音声認識の処理時間が短縮さ

れ、しいては音声対話のレスポンスが改善される。

【0014】

【発明の実施の形態】以下に、図面を参照して本発明の実施形態を説明する。図1は、本発明の1実施形態にかかる音声対話システム10の基本構成を示すブロック図である。このシステムは、カーナビゲーションシステムに音声による案内機能を提供するためのシステムであって、映像出力装置としてのディスプレイ12、音声出力装置としてのスピーカ14、音声入力装置としてのマイク16、音声対話開始スイッチ18及び処理装置20を備えている。

【0015】処理装置20は、ハードウェアとしては、周知のように、中央処理装置(CPU)、主記憶装置等からなるものである。また、処理装置20は、主記憶装置上で走行せしめられるソフトウェアとして、図1に示すように、ディスプレイ12に画像を表示する処理を行う画像表示エンジン202、マイク16から入力される音声のパターンを認識するための処理を行う音声認識エンジン204、およびスピーカ14から出力されるべき音声を電子的に合成する処理を行う音声合成エンジン206を備えている。

【0016】さらに、処理装置20は、ソフトウェアとして、図1に示すように、アプリケーションプログラム208の指示を受けて上述の画像表示エンジン、音声認識エンジン及び音声合成エンジンを制御する対話処理専用エンジン210を備えている。また、処理装置20は、上述の音声認識エンジン204、音声合成エンジン206及び対話処理専用エンジン210によって参照される対話用データベース212と主に音声認識エンジン204によって参照される音声認識辞書214を備える。

【0017】なお、当然のことながら、ハードウェア及びソフトウェアを総合的に管理及び制御するオペレーティングシステム(OS)216も処理装置20において走行する。また、上述の対話用データベース212および音声認識辞書214は、処理装置20内に組み込むことなく、CD-ROM、DVD-ROM、メモリカード等の外部メモリ上に構成されてもよい事は勿論である。

【0018】音声認識エンジン204は、図2に示す様に、音声認識のための処理プログラム(図示せず)と、カーナビゲーションシステムの操作に用いられる基本用語のための辞書204a、さらに音声認識辞書214から所定の部分を抽出して設定するRAM領域204bを有している。基本用語辞書204aには、対話に必要な基本的な用語、例えば「はい」、「いいえ」、「戻る」、「行く」などの言葉と、都道府県名、および頻繁に使用される大都市名、デパート、レストランなどの施設名などが、その音声パターンと共に記憶されている。尚、基本用語辞書204aは常時RAM領域204bに設定され、その語は音声認識可能となっている。また都

道府県名及び大都市名は基本用語辞書204aから外し、地名に関する辞書の最上位の階層として扱うことも可能である。

【0019】また音声認識エンジンは、探索対象単語属性リスト、対話状態とその対話に使用する辞書の種類などを記載したテーブルが保存される領域204c、探索対象単語属性リスト中にある単語属性が認識された場合それを保存するための保存領域204dなどが設けられているが、これらについては、各実施形態の説明の項で詳細に述べる。

【0020】音声認識辞書214には、図3に示す様に、住所の辞書、施設名称の辞書などが記憶されている。住所の辞書は、本発明では、図4に示す様に、概念的に上下関係を有する属性の単語を階層状に配置した階層構造を有すると共に、上位の階層の各単語をノードとして、このノードから関連する下位の階層の単語を分岐させた構造、即ちツリー構造を有している。

【0021】図示の例では、最上位の階層を「府県名」を属性に持つ単語で構成し、2段目の階層を「市名」を属性に持つ単語で構成し、3段目の階層を「区名」を属性に持つ単語で構成している。さらに、最上位の階層の府県名をそれぞれノードとし、各府県中の都市名を分岐させている。同様に、都市名の階層に属する各単語をノードとし、区名を分岐させている。これにより図4に示すような、ツリー状の3階層文法のイメージが構成される。

【0022】このような文法構造を有する住所の認識辞書を記憶する音声認識辞書214に対して、音声認識エンジン204は、前回のユーザ発声による音声認識の結果を参照して、特定のノード以下の文法を辞書214内から抽出し、これを自身のRAM領域204bに設定する。音声認識エンジン204はこの様にして設定された辞書を用いて次のユーザ発声を認識する。従って、この音声認識に当たっては、前回の認識結果によって規定されたノード以降の単語属性のデータ以外は音声認識エンジン204に設定されず、例えば兵庫県内の市名を発声したのにそれを大阪府内の市名と誤認識するような事態の発生は無い。従って、音声認識にあたって、認識率が向上する。

【0023】以下に、上記基本構造を有する音声対話システムの各種の実施形態の動作を、フローチャートを示して詳細に説明する。

【0024】図5は、本発明の第1の実施形態にかかる音声対話システムの動作説明のためのフローチャートである。また、図6はこのフローチャートの説明のための、音声対話システムとユーザとの対話例（対話例1）を示す。

【0025】システムが駆動設定されると、対話処理用エンジン210が作動して目的地設定のための最初の質問「どこへ行きたいですか」が作成され、この質問が音

声合成エンジン206によって音声に変換され、スピーカ14から出力される。この質問に対してユーザが「兵庫県」と答えると、音声認識エンジン204が自身の持っている基本辞書204aを参照してユーザ発声を認識する（図5のステップS1）。

【0026】次に、音声認識エンジン204は、最初の認識結果である単語属性「兵庫県」を、音声認識辞書204の住所辞書中で探索し（ステップS2）、探索された単語属性以降の辞書データを抽出して音声認識エンジン204のRAM領域204bに設定する（ステップS3）。具体的には、図4のノード10以下の辞書データが領域204bに設定される。

【0027】システムは次に「兵庫県のどこですか」を質問する。これに対してユーザが「神戸市」と答えると、音声認識エンジン204はRAM領域204bに設定された辞書データを用いてこれを認識する（ステップS4）。次のステップS5において対話が終了していないことを確認すると、音声認識エンジン204は再びステップS2に戻って、認識結果の単語属性である「神戸市」を図4の辞書中で探索し、ステップS3においてノード20以下の辞書データをRAM領域204bに設定する。

【0028】システムは次に「神戸市のどこですか」を質問し、ユーザはこれに対して「灘区」と答える。このユーザ発声は、RAM領域204bに設定されたノード20以下の辞書データを用いて認識される。これらの操作を対話終了、即ち目的地の発声が終了するまで繰り返すことにより、目的地の設定が完了する。

【0029】上述の実施形態では、「神戸市」の音声認識に当たって、音声認識エンジン204のRAM領域204bには、兵庫県以下の各都市名の辞書しか設定されないの、神戸市を間違えて他府県の都市名と誤認識する事は無い。同様に、「灘区」の認識に当たっても、RAM領域204bには、神戸市内の区名しか設定されないの、これを他市の区名と誤認識する可能性が排除される。

【0030】そのため、システムとユーザ間で、例えば、システム「どこに行きますか」、ユーザ「大阪府」、システム「大阪府のどこですか」、ユーザ「堺市」、システム「明石市のどこですか」（ユーザ発声「堺市」を「明石市」と誤認識）のような不自然な会話の発生が防止される。また、各認識に当たって、認識対象となる単語が絞り込まれるので、認識の速度が大幅に向上する。

【0031】図7は、本発明の第2の実施形態にかかる音声対話システムの動作説明のためのフローチャートであり、図8はこのフローチャートの説明に使用する対話例（対話例2）を示す。

【0032】この実施形態では、システム側の質問「どこへ行きたいですか」に対して、ユーザが「兵庫県神戸

市」と2個の単語で答えた場合に対処し得る構成を有している。

【0033】まず、ユーザの最初の発声「兵庫県神戸市」を、例えば基本辞書204aを用いて認識する(ステップS10)。次に音声認識エンジン204は、認識結果の単語属性「兵庫県」について、これを認識辞書214中で探索し(ステップS11)、探索された単語属性以下の辞書、即ち図4のノード10以下の辞書を抽出する(ステップS12)。次に、未探索の認識結果が存在するか否か判断し(ステップS13)、存在する場合、ステップS11に戻って未探索の認識結果に対して辞書中で探索を行なう。図8の場合は、未探索の認識結果「神戸市」が存在するのでこれをノード10以下の辞書中から探索し、ノード20以下の辞書を抽出する。

【0034】この操作を、未探索の認識結果が無くなるまで繰り返し、無くなると、ステップS14において最終的に抽出された辞書のデータを音声認識エンジン204のRAM領域204bに設定する。この後、次のユーザ発声「灘区」を、設定された辞書データを用いて認識する(ステップS15)。

【0035】以上の操作を対話終了まで繰り返す(ステップS16)事により、システムへの目的地の設定が完了する。

【0036】この実施形態では、ユーザによる複数の単語属性の発声に対して、音声認識エンジンが単語属性毎に繰り返して設定する辞書を絞り込むので、第1の実施形態の場合と同様、目的地の音声入力に当たって誤認識の発生が防止され、かつ高い応答速度を得ることができる。

【0037】図9は、本発明の第3の実施形態を説明するための図であって、具体的には音声認識エンジン206内に予め設定される探索対象単語属性リストを示す。この実施形態は、例えば、「大阪府に行く」と言うようなユーザの発声に対して、最初の音声認識で「大阪」、「に」、「行く」と言う3つの単語属性からなる認識結果を得ている場合に対処するものである。

【0038】本実施形態では、認識された各単語属性について、それぞれが探索対象単語属性リスト中に存在するか否かをまず探索する。図示の例では、「大阪府」のみがリスト中に存在するため、「大阪府」と言う単語属性に対して図5のフローチャートに示す処理を行なう。これによって、次のユーザ発声で大阪府の市区町村以外の単語属性を持つ結果が認識されることを防止することが出来る。

【0039】図10は、本発明にかかる第4の実施形態の構成を説明するための対話例(対話例3)を示す図である。この実施形態では、過去に認識された認識結果について、図9に示すような探索対象単語属性リストに属する単語属性があった場合、これを認識エンジン204の保存領域204dに保存しておき、その後、ユーザ発

声の認識のための辞書データを作成するに当たって、前回の認識結果に加えて保存された結果を使用することを特徴とする。

【0040】即ち、図10に示す様に、システム側の質問「どこに行きますか」に対するユーザの回答「兵庫県」が認識されると、図5または図7のフローチャートに示す様に、音声認識エンジン204中のRAM領域206bには「兵庫県」をノードとするそれ以降の辞書データが設定される。認識された単語属性「兵庫県」は、図9に示す探索対象単語属性リスト中に含まれているので、この単語属性「兵庫県」は、音声認識エンジン204の保存領域204d中に保存される。

【0041】システムは、次に「兵庫県のどこですか」を質問する。ユーザはこれに対して「ラーメン屋」と答える。「ラーメン屋」と言う単語属性は、音声認識エンジン204中の基本辞書における施設の種類の辞書中に含まれているので、その波形データを使用して「ラーメン屋」が認識される。

【0042】「ラーメン屋」が認識されると、音声認識エンジンは、音声認識辞書214中の施設名の辞書に含まれている、ラーメン屋の辞書を抽出し、さらに保存領域204d中に保存された単語属性「兵庫県」を用いてラーメン屋の辞書をさらに絞り込む。これによって、兵庫県内のラーメン屋の辞書が抽出され、これが音声認識エンジン204のRAM領域204bに設定される。次のユーザ発声「〇〇ラーメン」は、この様にして設定された辞書データを用いて認識される。

【0043】なお、施設名の単語属性には、それぞれの住所がデータとして添付されているので、このデータを探索することにより、特定の地域のラーメン屋を特定することができる。

【0044】図11は、本発明にかかる第5の実施形態の構成を説明するための対話例(対話例4)を示す図である。この実施形態では、過去に認識された認識結果について、図9に示すような探索対象単語属性リストに属する単語属性があった場合、これを認識エンジン204の保存領域204dに保存しておき、その後、システムとユーザの対話の過程で認識間違いがあった場合、保存領域204dに保存されている単語属性を用いて辞書の絞り込みを行なうことを特徴としている。

【0045】即ち、図11に示す様に、システム側の質問「どこに行きますか」に対するユーザの回答「兵庫県」が認識されると、図5または図7のフローチャートに示す様に、音声認識エンジン204中のRAM領域206bには「兵庫県」をノードとするそれ以降の辞書データが設定される。認識された単語属性「兵庫県」は、図9に示す探索対象単語属性リスト中に含まれているので、音声認識エンジン204の保存領域204d中に保存される。

【0046】システムは、次に「兵庫県のどこですか」

を質問する。ユーザはこれに対して「神戸市」と答えるが、システムがこれを「明石市」と誤認識し、次の質問「明石市のどこですか」を発話する。ユーザはシステムのこの質問によって前回の自身の回答が誤認識されたことを知って、「ちがう」と回答する。システムが基本辞書を用いてこの発声を認識すると、保存領域204dを探索して保存されている単語属性「兵庫県」を見出す。次に、この単語属性を用いて辞書の絞り込みを行なうと共に、質問「兵庫県のどこですか」を再度作成する。

【0047】この操作は、図11において二重線矢印で示すように、認識を誤った場所まで対話が戻ることを示している。システムの再度の質問以降は、図の二重線矢印で示す様に、上記各実施形態の処理が実行される。

【0048】図12～14は、本発明の第6の実施形態を説明するための図であって、図12はこの実施形態に適用される対話例（対話例5）を示し、図13、14は本実施形態における音声認識エンジン204と音声認識辞書214の構造を概念的に示す図である。

【0049】本実施形態では、ユーザとシステムの対話状態（図12参照）に応じて使用する辞書の種類を予め設定しておき、各対話状態に達したとき、予め設定された種類の辞書のデータを音声認識辞書から抜き出して、音声認識エンジン204のRAM領域204bに設定するようにしたことを特徴とする。

【0050】例えば、対話状態1～4に対して図13に示すような対応テーブル204eが予め形成されていると、対話状態4では、該対応テーブルに従って施設名称の辞書のデータを音声認識辞書214から取り出して来てRAM領域204bに設定し、この辞書を用いて次のユーザ発声の音声認識を実行する。

【0051】あるいは、対話状態2の場合は、図14に示す様に、対応テーブル204eに従って住所の辞書と施設の種類の辞書を音声認識辞書214から取り出し、それらを共にRAM領域204bに設定する。これにより、ユーザの次の発声が「ラーメン屋」であっても容易に音声認識される。

【0052】本実施形態によれば、対話の各状態において、ユーザが発生すると思われる発声パターンを有する辞書のみを音声認識辞書から取り出し、音声認識エンジンに設定することで、該エンジンにかかる負荷（処理量）を低減することができる。その結果、音声認識の応答時間が短縮でき、音声対話のレスポンスが改善される。

【0053】本発明のさらに他の実施形態として、ある状態で、既に得られている単語属性以外の、対話タスクを達成させるに必要な単語属性を含む音声認識用辞書を自動的に選択し、音声認識エンジンに設定する事も可能である。即ち、目的地設定の対話で、既に「県名」の属性を持つ認識結果が得られている場合、タスク達成のために他に必要な属性は、「市名」、「町名」、「番地」

などであり、これらの単語属性を含む辞書を、複数の辞書から探索し、適合する辞書を組み合わせて音声認識エンジンに設定する。

【0054】これによって、第6の実施形態のように対応テーブルを持つことなく、この場合の実施形態と同じ効果を得ることができ、処理量（データ量）の削減が期待できる。

【0055】

【発明の効果】以上、各実施形態を挙げて説明したように、本発明の音声対話システムによれば、既に得られている音声認識の結果に関連する辞書データを音声認識辞書から取り出して音声認識エンジンに設定し、次のユーザ発声の音声認識を行なうので、音声認識の誤認率が低下する。同時に音声認識エンジンに設定されるデータ量が減少するので認識エンジンの負荷が減少し、認識速度が速くなる。その結果、対話の応答速度が向上する。

【図面の簡単な説明】

【図1】本発明の1実施形態にかかる音声対話システムの構成を示すブロック図。

【図2】図1に示す音声認識エンジンの構成を示す図。

【図3】図1に示す音声認識辞書の構成を示す図。

【図4】ツリー状階層構造を有する住所辞書の構成を示す図。

【図5】本発明の第1の実施形態にかかる音声対話システムの動作説明のためのフローチャート。

【図6】図5のフローチャートの説明に提供する第1の対話例。

【図7】本発明の第2の実施形態にかかる音声対話システムの動作説明のためのフローチャート。

【図8】図7のフローチャートの説明に提供する第2の対話例。

【図9】探索対象単語属性リストの一例を示す図。

【図10】本発明の第3の実施形態の説明に供する第3の対話例。

【図11】本発明の第4の実施形態の説明に供する第4の対話例。

【図12】本発明の第5の実施形態の説明に供する第5の対話例。

【図13】本発明の第6の実施形態の説明に供する図。

【図14】本発明の第6の実施形態の説明に供する図。

【符号の説明】

10…音声対話システム

12…ディスプレイ

14…スピーカ

16…マイク

18…スイッチ

20…処理装置

202…画像表示エンジン

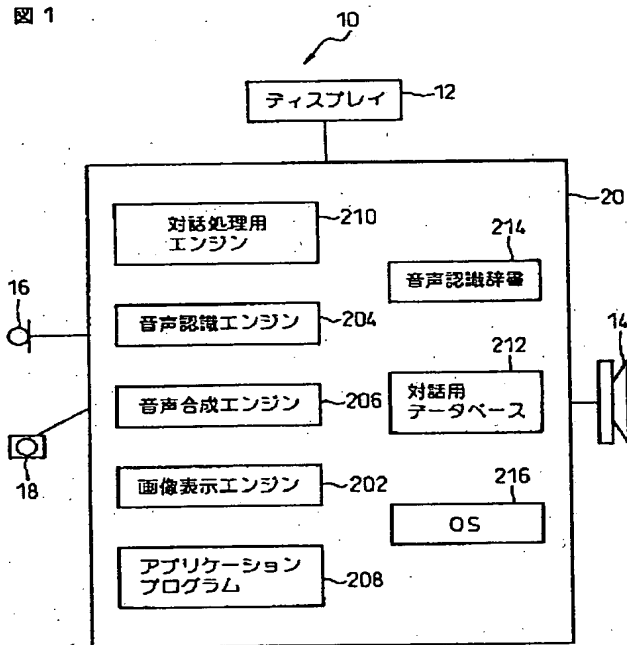
204…音声認識エンジン

206…音声合成エンジン

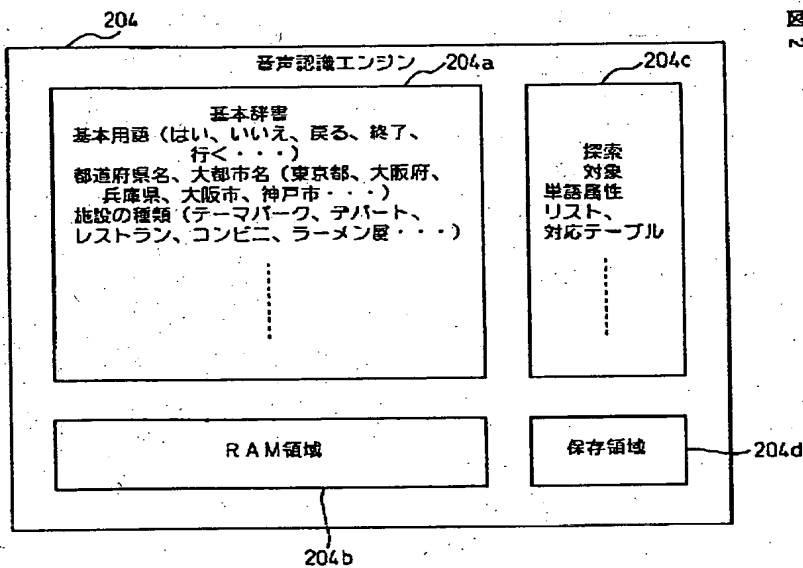
11  
208…アプリケーションプログラム  
210…対話処理用エンジン

12  
212…対話用データベース  
214…音声認識辞書

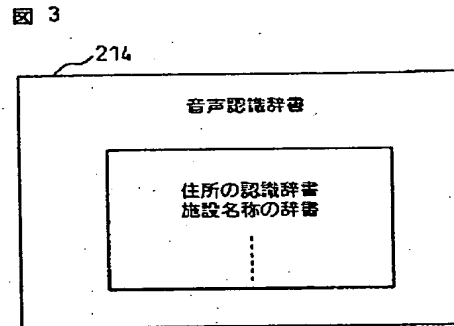
【図1】



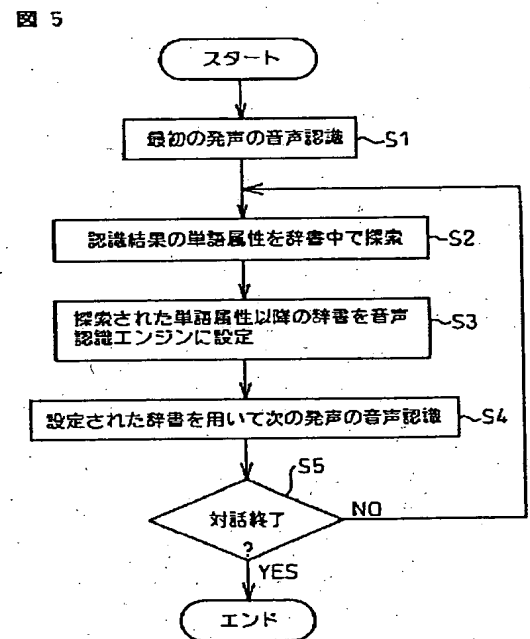
【図2】



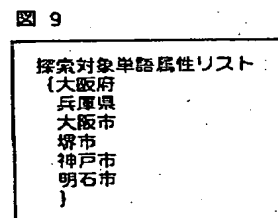
【図3】



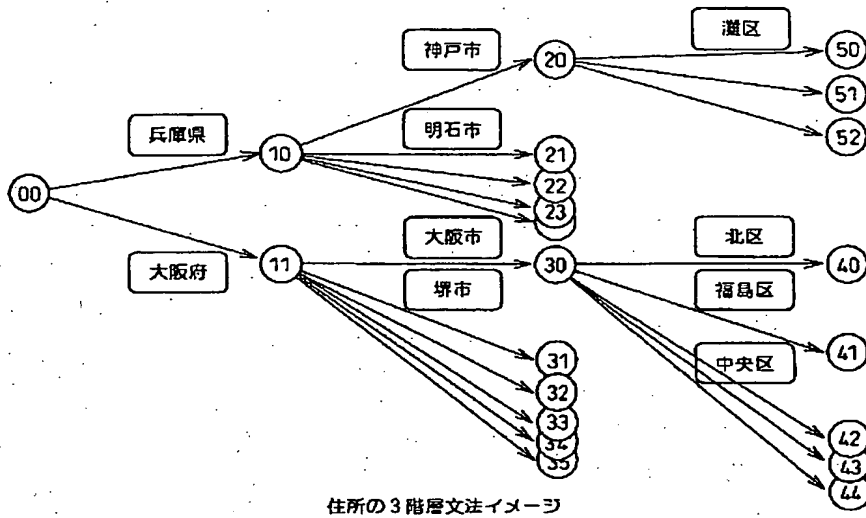
【図5】



【図9】



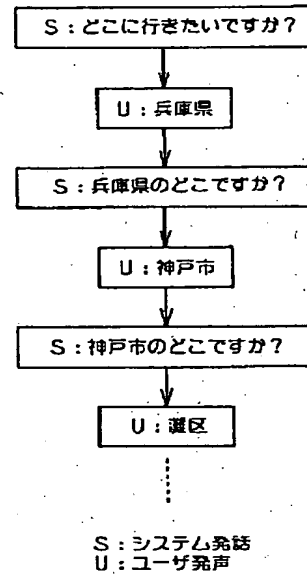
【図4】



【図6】

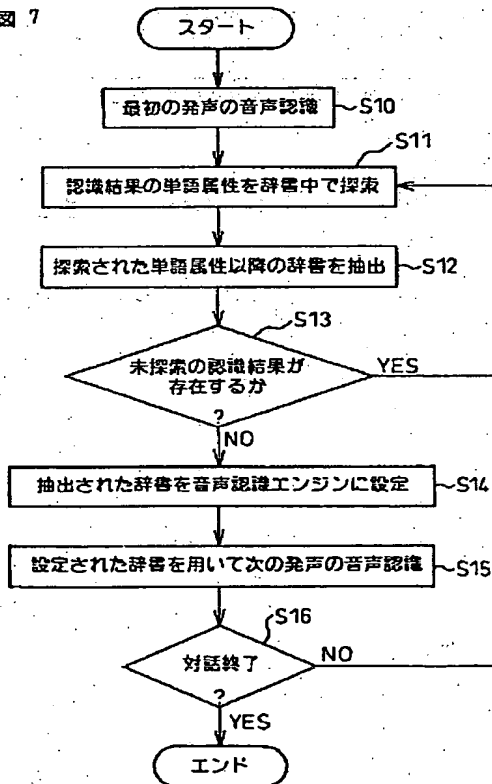
図 図 6

対話例1



【図7】

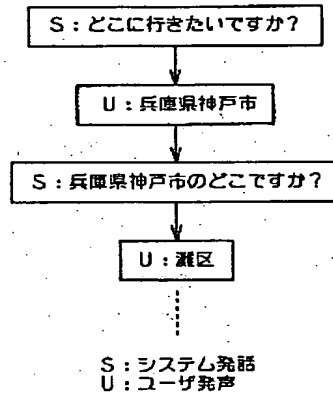
図 7



【図8】

図 8

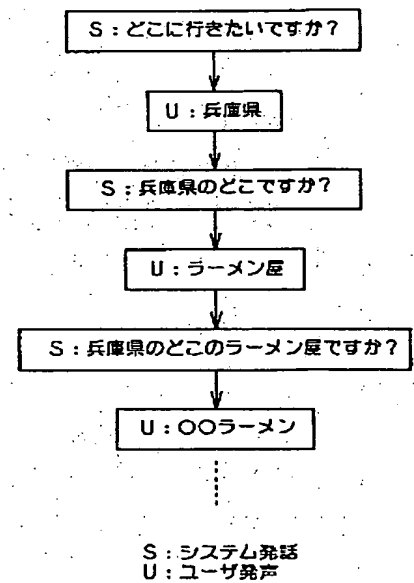
対話例2



【図10】

図 10

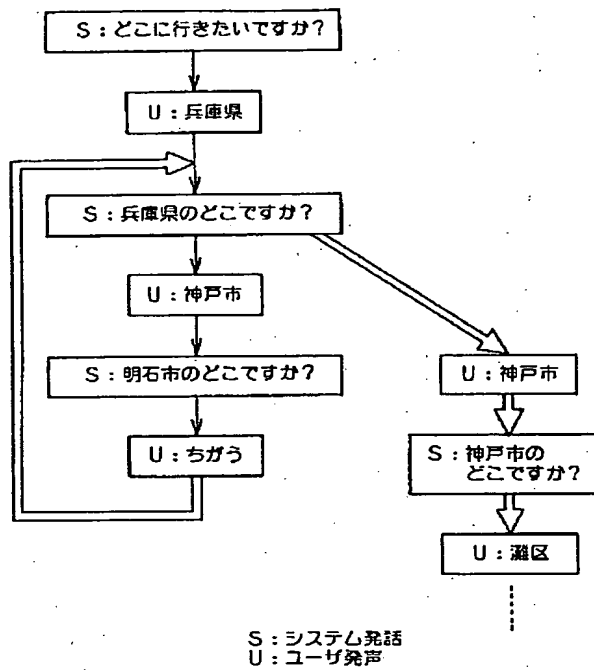
対話例3



【図11】

図 11

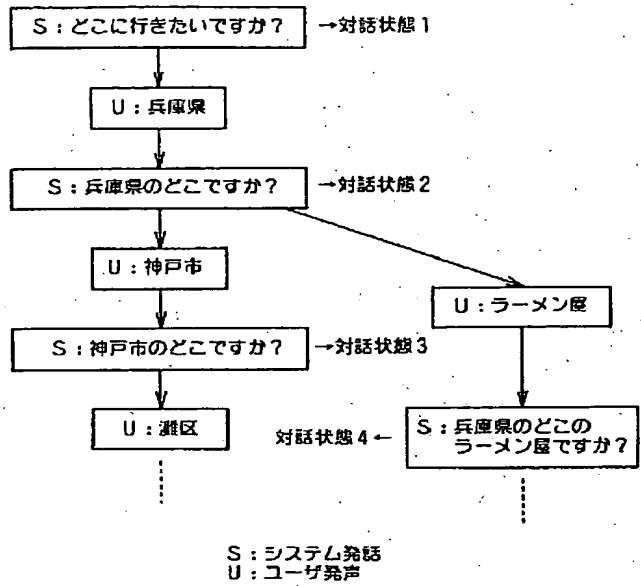
対話例4



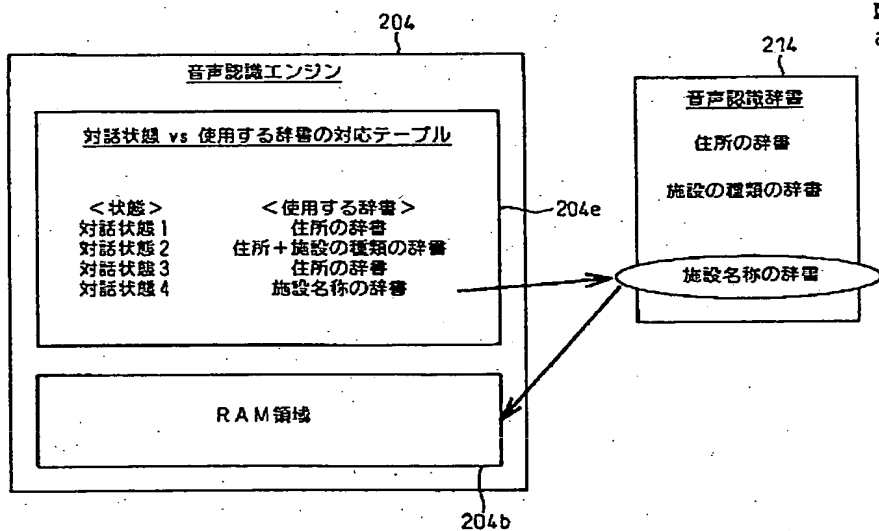
【図12】

図 12

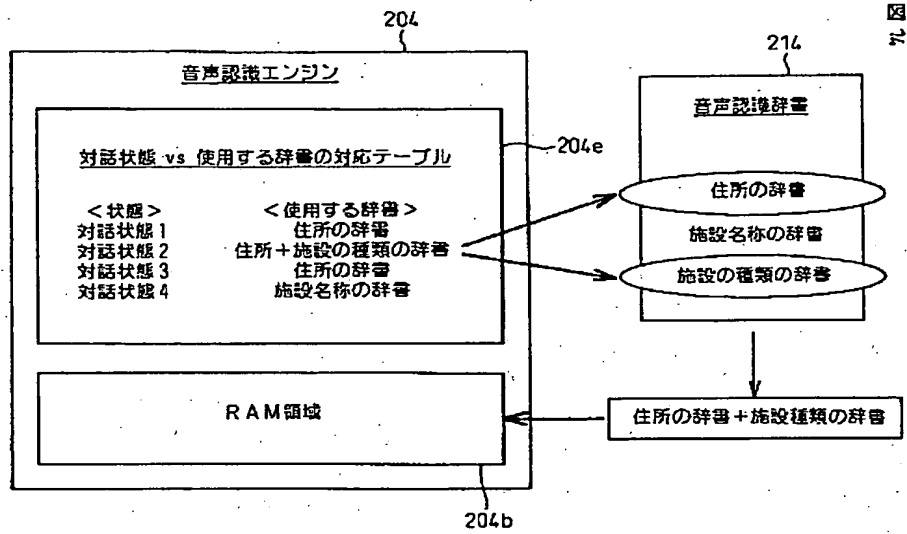
対話例5



【図13】



【図14】



フロントページの続き

(51) Int. Cl.<sup>7</sup>

G10L 15/18

15/28

識別記号

FI

G10L 3/00

キーワード(参考)

537F

521W

This Page is inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images problems checked, please do not report the problems to the IFW Image Problem Mailbox**